

Rec'd PCT/PTO 04 APR 2005

10/530253

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.

Application Serial Number: 10/S30/253
Source: PCT
Date Processed by STIC: 4-13-05

ENTERED



PCT

RAW SEQUENCE LISTING

DATE: 04/13/2005

PATENT APPLICATION: US/10/530,253

TIME: 09:34:29

Input Set : A:\00400769.txt

Output Set: N:\CRF4\04132005\J530253.raw

3 <110> APPLICANT: Cassetti, Maria C.
 4 Smith, Larry
 5 Jeffrey K. Pullen
 6 Susan P. McElhiney
 8 <120> TITLE OF INVENTION: HUMAN PAPILLOMAVIRUS POLYPEPTIDES AND IMMUNOGENIC
 COMPOSITIONS
 10 <130> FILE REFERENCE: 00630/100M137-US2
 C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/530,253
 C--> 12 <141> CURRENT FILING DATE: 2005-04-04
 12 <150> PRIOR APPLICATION NUMBER: PCT/US2003/031726
 13 <151> PRIOR FILING DATE: 2003-10-02
 15 <150> PRIOR APPLICATION NUMBER: US 60/415,929
 16 <151> PRIOR FILING DATE: 2002-10-03
 18 <160> NUMBER OF SEQ ID NOS: 65
 20 <170> SOFTWARE: PatentIn version 3.1
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 248
 24 <212> TYPE: PRT
 25 <213> ORGANISM: Human papillomavirus type 16
 27 <400> SEQUENCE: 1
 29 Met Phe Gln Asp Pro Gln Glu Arg Pro Arg Lys Leu Pro Gln Leu Cys
 30 1 5 10 15
 33 Thr Glu Leu Gln Thr Thr Ile His Asp Ile Ile Leu Glu Cys Val Tyr
 34 20 25 30
 37 Cys Lys Gln Gln Leu Leu Arg Arg Glu Val Tyr Asp Phe Ala Phe Arg
 38 35 40 45
 41 Asp Leu Cys Ile Val Tyr Arg Asp Gly Asn Pro Tyr Ala Val Cys Asp
 42 50 55 60
 45 Lys Cys Leu Lys Phe Tyr Ser Lys Ile Ser Glu Tyr Arg His Tyr Cys
 46 65 70 75 80
 49 Tyr Ser Val Tyr Gly Thr Thr Leu Glu Gln Gln Tyr Asn Lys Pro Leu
 50 85 90 95
 53 Cys Asp Leu Leu Ile Arg Cys Ile Asn Cys Gln Lys Pro Leu Cys Pro
 54 100 105 110
 57 Glu Glu Lys Gln Arg His Leu Asp Lys Lys Gln Arg Phe His Asn Ile
 58 115 120 125
 61 Arg Gly Arg Trp Thr Gly Arg Cys Met Ser Cys Cys Arg Ser Ser Arg
 62 130 135 140
 65 Thr Arg Arg Glu Thr Gln Leu His Gly Asp Thr Pro Thr Leu His Glu
 66 145 150 155 160
 69 Tyr Met Leu Asp Leu Gln Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu
 70 165 170 175
 73 Gln Leu Asn Asp Ser Ser Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala
 74 180 185 190

(ps.6)

RAW SEQUENCE LISTING

DATE: 04/13/2005

PATENT APPLICATION: US/10/530,253

TIME: 09:34:29

Input Set : A:\00400769.txt

Output Set: N:\CRF4\04132005\J530253.raw

```

77 Gly Gln Ala Glu Pro Asp Arg Ala His Tyr Asn Ile Val Thr Phe Cys
78      195      200      205
81 Cys Lys Cys Asp Ser Thr Leu Arg Leu Cys Val Gln Ser Thr His Val
82      210      215      220
85 Asp Ile Arg Thr Leu Glu Asp Leu Leu Met Gly Thr Leu Gly Ile Val
86 225      230      235      240
89 Cys Pro Ile Cys Ser Gln Lys Pro
90      245
93 <210> SEQ ID NO: 2
94 <211> LENGTH: 747
95 <212> TYPE: DNA
96 <213> ORGANISM: Human papillomavirus type 16
98 <400> SEQUENCE: 2
99 atgtttcagg acccacagga gcgacccaga aagttaccac agttatgcac agagctgcaa      60
101 acaactatac atgatataat attagaatgt gtgtactgca agcaacagtt actgcgacgt      120
103 gaggtatatg acttttgcttt tcgggattta tgcatagtat atagagatgg gaatccatat      180
105 gctgtatgtg ataaatgttt aaagttttat tctaaaatta gtgagtatag acattattgt      240
107 tatagttgtt atggaacaac attagaacag caatacaaca aaccgttgtg tgatttgta      300
109 attaggtgta ttaactgtca aaagccactg tgcctgaag aaaagcaaag acatctggac      360
111 aaaaagcaaa gattccataa tataaggggt cggaggaccg gtcgatgtat gtcttggtgc      420
113 agatcatcaa gaacacgtag agaaaccag ctgcattgag atacacctac attgcatgaa      480
115 tatatgttag atttgcaacc agagacaact gatctctact gttatgagca attaatgac      540
117 agctcagagg aggaggatga aatagatggt ccagctggac aagcagaacc ggacagagcc      600
119 cattacaata ttgtaacctt ttgttgcaag tgtgactcta cgcttcggtt gtgcgtacaa      660
121 agcacacacg tagacattcg tactttggaa gacctgttaa tgggcacact aggaattgtg      720
123 tgccccatct gttctcagaa accataa      747
126 <210> SEQ ID NO: 3
127 <211> LENGTH: 248
128 <212> TYPE: PRT
129 <213> ORGANISM: Human papillomavirus type 16
131 <400> SEQUENCE: 3
133 Met Phe Gln Asp Pro Gln Glu Arg Pro Arg Lys Leu Pro Gln Leu Cys
134 1      5      10      15
137 Thr Glu Leu Gln Thr Thr Ile His Asp Ile Ile Leu Glu Cys Val Tyr
138      20      25      30
141 Cys Lys Gln Gln Leu Leu Arg Arg Glu Val Tyr Asp Phe Ala Phe Arg
142      35      40      45
145 Asp Leu Cys Ile Val Tyr Arg Asp Gly Asn Pro Tyr Ala Val Gly Asp
146      50      55      60
149 Lys Cys Leu Lys Phe Tyr Ser Lys Ile Ser Glu Tyr Arg His Tyr Cys
150 65      70      75      80
153 Tyr Ser Val Tyr Gly Thr Thr Leu Glu Gln Gln Tyr Asn Lys Pro Leu
154      85      90      95
157 Cys Asp Leu Leu Ile Arg Cys Ile Asn Gly Gln Lys Pro Leu Cys Pro
158      100      105      110
161 Glu Glu Lys Gln Arg His Leu Asp Lys Lys Gln Arg Phe His Asn Ile
162      115      120      125
165 Arg Gly Arg Trp Thr Gly Arg Cys Met Ser Cys Cys Arg Ser Ser Arg
166      130      135      140

```

RAW SEQUENCE LISTING

DATE: 04/13/2005

PATENT APPLICATION: US/10/530,253

TIME: 09:34:29

Input Set : A:\00400769.txt

Output Set: N:\CRF4\04132005\J530253.raw

```

169 Thr Arg Arg Glu Thr Gln Leu His Gly Asp Thr Pro Thr Leu His Glu
170 145 150 155 160
173 Tyr Met Leu Asp Leu Gln Pro Glu Thr Thr Asp Leu Tyr Gly Tyr Gly
174 165 170 175
177 Gln Leu Asn Asp Ser Ser Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala
178 180 185 190
181 Gly Gln Ala Glu Pro Asp Arg Ala His Tyr Asn Ile Val Thr Phe Cys
182 195 200 205
185 Cys Lys Cys Asp Ser Thr Leu Arg Leu Cys Val Gln Ser Thr His Val
186 210 215 220
189 Asp Ile Arg Thr Leu Glu Asp Leu Leu Met Gly Thr Leu Gly Ile Val
190 225 230 235 240
193 Cys Pro Ile Cys Ser Gln Lys Pro
194 245
197 <210> SEQ ID NO: 4
198 <211> LENGTH: 747
199 <212> TYPE: DNA
200 <213> ORGANISM: Human papillomavirus type 16
202 <400> SEQUENCE: 4
203 atgtttcagg acccacagga gcgaccacaga aagttaccac agttatgcac agagctgcaa 60
205 acaactatac atgatataat attagaatgt gtgtactgca agcaacagtt actgcgacgt 120
207 gaggtatatg actttgcttt tcgggattta tgcatagtat atagagatgg gaatccatat 180
209 gctgtagggtg ataaatgttt aaagttttat tctaaaatta gtgagtatag acattattgt 240
211 tatagtgtgt atggaacaac attagaacag caatacaaca aaccgttgtg tgatttggtta 300
213 attaggtgta ttaacgggtca aaagccactg tgtcctgaag aaaagcaaag acatctggac 360
215 aaaaagcaaa gattccataa tataaggggt cggtggaccg gtcgatgtat gtcttggtgc 420
217 agatcatcaa gaacacgtag agaaaccag ctgcatggag atacacctac attgcatgaa 480
219 tatatgtagg atttgcaacc agagacaact gatctctacg gttatgggca attaaatgac 540
221 agctcagagg aggaggatga aatagatggt ccagctggac aagcagaacc ggacagagcc 600
223 cattacaata ttgtaacctt ttgttgcaag tgtgactcta cgcttcgggt gtgcgtacaa 660
225 agcacacacg tagacattcg tactttggaa gacctgttaa tgggcacact aggaattgtg 720
227 tgccccatct gttctcagaa accataa 747
230 <210> SEQ ID NO: 5
231 <211> LENGTH: 248
232 <212> TYPE: PRT
233 <213> ORGANISM: Human papillomavirus type 16
235 <400> SEQUENCE: 5
237 Met Phe Gln Asp Pro Gln Glu Arg Pro Arg Lys Leu Pro Gln Leu Cys
238 1 5 10 15
241 Thr Glu Leu Gln Thr Thr Ile His Asp Ile Ile Leu Glu Cys Val Tyr
242 20 25 30
245 Cys Lys Gln Gln Leu Leu Arg Arg Glu Val Tyr Asp Phe Ala Phe Arg
246 35 40 45
249 Asp Leu Cys Ile Val Tyr Arg Asp Gly Asn Pro Tyr Ala Val Gly Asp
250 50 55 60
253 Lys Cys Leu Lys Phe Tyr Ser Lys Ile Ser Glu Tyr Arg His Tyr Cys
254 65 70 75 80
257 Tyr Ser Val Tyr Gly Thr Thr Leu Glu Gln Gln Tyr Asn Lys Pro Leu
258 85 90 95

```

RAW SEQUENCE LISTING

DATE: 04/13/2005

PATENT APPLICATION: US/10/530,253

TIME: 09:34:29

Input Set : A:\00400769.txt

Output Set: N:\CRF4\04132005\J530253.raw

```

261 Cys Asp Leu Leu Ile Arg Cys Ile Asn Gly Gln Lys Pro Leu Cys Pro
262           100           105           110
265 Glu Glu Lys Gln Arg His Leu Asp Lys Lys Gln Arg Phe His Asn Ile
266           115           120           125
269 Arg Gly Arg Trp Thr Gly Arg Cys Met Ser Cys Cys Arg Ser Ser Arg
270           130           135           140
273 Thr Arg Arg Glu Thr Gln Leu His Gly Asp Thr Pro Thr Leu His Glu
274 145           150           155           160
277 Tyr Met Leu Asp Leu Gln Pro Glu Thr Thr Asp Leu Tyr Gly Tyr Gly
278           165           170           175
281 Gln Leu Asn Asp Ser Ser Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala
282           180           185           190
285 Gly Gln Ala Glu Pro Asp Arg Ala His Tyr Asn Ile Val Thr Phe Cys
286           195           200           205
289 Cys Lys Cys Asp Ser Thr Leu Arg Leu Cys Val Gln Ser Thr His Val
290           210           215           220
293 Asp Ile Arg Thr Leu Glu Asp Leu Leu Met Gly Thr Leu Gly Ile Val
294 225           230           235           240
297 Gly Pro Ile Cys Ser Gln Lys Pro
298           245

```

301 <210> SEQ ID NO: 6

302 <211> LENGTH: 747

303 <212> TYPE: DNA

304 <213> ORGANISM: Human papillomavirus type 16

306 <400> SEQUENCE: 6

```

307 atgtttcagg acccacagga gcgaccaga aagttaccac agttatgcac agagctgcaa      60
309 acaactatac atgatataat attagaatgt gtgtactgca agcaacagtt actgcgacgt      120
311 gaggtatatg actttgcttt tcgggattta tgcatagtat atagagatgg gaatccatat      180
313 gctgtagggt ataaatgttt aaagttttat tctaaaatta gtgagtatag acattattgt      240
315 tatagtgtgt atggaacaac attagaacag caatacaaca aaccgttgtg tgatttgta      300
317 attagtgta ttaacgggtca aaagccactg tgtcctgaag aaaagcaaag acatctggac      360
319 aaaaagcaaa gattccataa tataaggggt cggtggaccg gtcgatgtat gtcttgttgc      420
321 agatcatcaa gaacacgtag agaaaccag ctgcattggag atacacctac attgcatgaa      480
323 tatatgttag atttgaacc agagacaact gatctctacg gttatgggca attaatgac      540
325 agctcagagg aggaggatga aatagatggt ccagctggac aagcagaacc ggacagagcc      600
327 cattacaata ttgtaacctt ttgttgcaag tgtgactcta cgcttcggtt gtgcgtacaa      660
329 agcacacacg tagacattcg tactttggaa gacctgttaa tgggcacact aggaattgtg      720
331 ggccccatct gttctcagaa accataa                                747

```

334 <210> SEQ ID NO: 7

335 <211> LENGTH: 248

336 <212> TYPE: PRT

337 <213> ORGANISM: Human papillomavirus type 16

339 <400> SEQUENCE: 7

```

341 Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp Leu Gln
342 1           5           10           15
345 Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu Gln Leu Asn Asp Ser Ser
346           20           25           30
349 Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala Glu Pro Asp
350           35           40           45

```

RAW SEQUENCE LISTING

DATE: 04/13/2005

PATENT APPLICATION: US/10/530,253

TIME: 09:34:29

Input Set : A:\00400769.txt

Output Set: N:\CRF4\04132005\J530253.raw

```

353 Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys Asp Ser Thr
354      50                      55                      60
357 Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg Thr Leu Glu
358 65                      70                      75                      80
361 Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile Cys Ser Gln
362      85                      90                      95
365 Lys Pro Phe Gln Asp Pro Gln Glu Arg Pro Arg Lys Leu Pro Gln Leu
366      100                     105                     110
369 Cys Thr Glu Leu Gln Thr Thr Ile His Asp Ile Ile Leu Glu Cys Val
370      115                     120                     125
373 Tyr Cys Lys Gln Gln Leu Leu Arg Arg Glu Val Tyr Asp Phe Ala Phe
374      130                     135                     140
377 Arg Asp Leu Cys Ile Val Tyr Arg Asp Gly Asn Pro Tyr Ala Val Cys
378 145                     150                     155                     160
381 Asp Lys Cys Leu Lys Phe Tyr Ser Lys Ile Ser Glu Tyr Arg His Tyr
382      165                     170                     175
385 Cys Tyr Ser Val Tyr Gly Thr Thr Leu Glu Gln Gln Tyr Asn Lys Pro
386      180                     185                     190
389 Leu Cys Asp Leu Leu Ile Arg Cys Ile Asn Cys Gln Lys Pro Leu Cys
390      195                     200                     205
393 Pro Glu Glu Lys Gln Arg His Leu Asp Lys Lys Gln Arg Phe His Asn
394      210                     215                     220
397 Ile Arg Gly Arg Trp Thr Gly Arg Cys Met Ser Cys Cys Arg Ser Ser
398 225                     230                     235                     240
401 Arg Thr Arg Arg Glu Thr Gln Leu
402      245

```

405 <210> SEQ ID NO: 8

406 <211> LENGTH: 747

407 <212> TYPE: DNA

408 <213> ORGANISM: Human papillomavirus type 16

410 <400> SEQUENCE: 8

```

411 atgcatggag atacacctac attgcatgaa tatatgttag atttgcaacc agagacaact      60
413 gatctctact gttatgagca attaaatgac agctcagagg aggaggatga aatagatggg      120
415 ccagctggac aagcagaacc ggacagagcc cattacaata ttgtaacctt ttgttgcaag      180
417 tgtgactcta cgcttcggtt gtgcgtacaa agcacacacg tagacattcg tactttggaa      240
419 gacctgttaa tgggcacact aggaattgtg tgccccatct gttctcagaa accatttcag      300
421 gaccacagc agcgaccacg aaagttacca cagttatgca cagagctgca aacaactata      360
423 catgatataa tattagaatg tgtgtactgc aagcaacagt tactgcgacg tgaggtatat      420
425 gactttgctt ttcgggattt atgcatagta tatagagatg ggaatccata tgctgtatgt      480
427 gataaatggt taaagtttta ttctaaaatt agtgagtata gacattattg ttatagtgtg      540
429 tatggaacaa cattagaaca gcaatacaac aaaccgttgt gtgatttgtt aattaggtgt      600
431 attaaactgtc aaaagccact gtgtcctgaa gaaaagcaaa gacatctgga caaaaagcaa      660
433 agattccata atataagggg tcggtggacc ggtcgatgta tgtcttggtg cagatcatca      720
435 agaacacgta gagaaaccca gctgtaa                                747

```

438 <210> SEQ ID NO: 9

439 <211> LENGTH: 248

440 <212> TYPE: PRT

441 <213> ORGANISM: Human papillomavirus type 16

443 <400> SEQUENCE: 9

RAW SEQUENCE LISTING ERROR SUMMARY

DATE: 04/13/2005

PATENT APPLICATION: US/10/530,253

TIME: 09:34:30

Input Set : A:\00400769.txt

Output Set: N:\CRF4\04132005\J530253.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:39; Xaa Pos. 1,2,3,8,12,15,19,20,22,23,24,25,29,30,31,37,42,46,50,53
Seq#:39; Xaa Pos. 59,60,62,64,66,67,70,76,78,80,82,83,88,92,93,95,97,99,106
Seq#:39; Xaa Pos. 107,110,115,118,121,123,125,131,133,137,139,140,143,144
Seq#:39; Xaa Pos. 145,147,148,149,150,151,152
Seq#:40; Xaa Pos. 2,4,5,9,10,11,16,19,20,21,24,30,33,34,36,37,38,39,40,41
Seq#:40; Xaa Pos. 42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,59,61,62
Seq#:40; Xaa Pos. 63,66,68,69,70,71,72,74,78,79,80,84,94,95,99,102,103,104

VERIFICATION SUMMARY

DATE: 04/13/2005

PATENT APPLICATION: US/10/530,253

TIME: 09:34:30

Input Set : A:\00400769.txt

Output Set: N:\CRF4\04132005\J530253.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application No
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:1725 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:39 after pos.:0
M:341 Repeated in SeqNo=39
L:1778 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:40 after pos.:0
M:341 Repeated in SeqNo=40